

Advanced and Intelligent Manufacturing in China

Li Li · Qingyun Yu · Kuo-Yi Lin ·  
Yumin Ma · Fei Qiao

---

# Data-Driven Scheduling of Semiconductor Manufacturing Systems



Chemical Industry Press Co., Ltd.



Springer

# **Advanced and Intelligent Manufacturing in China**

**Series Editor**

Jie Chen, Tongji University, Shanghai, China

This is a set of high-level and original academic monographs. This series focuses on the two fields of intelligent manufacturing and equipment, control and information technology, covering a range of core technologies such as Internet of Things, 3D printing, robotics, intelligent equipment, and epitomizing the achievements of technological development in China's manufacturing sector.

With Prof. Jie Chen, a member of the Chinese Academy of Engineering and a control engineering expert in China, as the Editorial in Chief, this series is organized and written by more than 30 young experts and scholars from more than 10 universities and institutes. It typically embodies the technological development achievements of China's manufacturing industry. It will promote the research and development and innovation of advanced intelligent manufacturing technologies, and promote the technological transformation and upgrading of the equipment manufacturing industry.

Li Li · Qingyun Yu · Kuo-Yi Lin · Yumin Ma ·  
Fei Qiao

# Data-Driven Scheduling of Semiconductor Manufacturing Systems



Chemical Industry Press Co., Ltd.



Springer

Li Li  
Tongji University  
Shanghai, China

Qingyun Yu  
Tongji University  
Shanghai, China

Kuo-Yi Lin  
Tongji University  
Shanghai, China

Yumin Ma  
Tongji University  
Shanghai, China

Fei Qiao  
Tongji University  
Shanghai, China

B&R Book Program

ISSN 2731-5983

ISSN 2731-5991 (electronic)

Advanced and Intelligent Manufacturing in China

ISBN 978-981-19-7587-5

ISBN 978-981-19-7588-2 (eBook)

<https://doi.org/10.1007/978-981-19-7588-2>

Jointly published with Chemical Industry Press, Beijing, China

The print edition is not for sale in China (Mainland). Customers from China (Mainland) please order the print book from: Chemical Industry Press.

© Chemical Industry Press 2023

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publishers, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publishers nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publishers remain neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.

The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

# Preface

Scheduling problem is ubiquitous in industrial engineering; its essence is to seek the maximization of system objectives through the reasonable allocation of limited resources. There are extensive and diverse objective contradictions between the limitation of resources (including material resources and time resources) and the pursuit of objectives (such as output, efficiency, speed, etc.). Therefore, scheduling problem has always been one of the research hotspots in academic and engineering technology circles.

Although there are many kinds of practical performance of scheduling problem, such as manufacturing system scheduling, traffic and transportation scheduling, personnel time scheduling, project schedule scheduling, manufacturing system scheduling is undoubtedly the most concerning and has been studied for the longest time among many scheduling problems. Manufacturing system scheduling is the central problem of enterprise production activities organization and management, which is an effective way to improve enterprise comprehensive benefits. It is of great significance to improve enterprise production management level, save cost, improve service quality, enhance enterprise competitiveness, accelerate the recovery of investment, and obtain higher economic benefits. With the help of advanced production planning and scheduling methods, enterprises can gain greater output and profit and higher rate of return on investment on the basis of no or less increase in investment.

Since Johnson published the first classic paper on production scheduling in 1954, the scheduling problem of manufacturing system has gone through the development process from simple to complex, such as single machine scheduling, multi-machine scheduling, flow-shop scheduling, job-shop scheduling, flexible manufacturing system (FMS) scheduling, and so on. During this period, the accumulation of a large number of related research work and achievements has further established the pivotal position of manufacturing system scheduling in the field of scheduling research. Many early researches in the scheduling field were developed under the promotion of the manufacturing industry, the practical problems arising from manufacturing continue to present new challenges. According to Lawler et al., with the continuous breakthrough of the four basic assumptions of the classical scheduling

problem (i.e., single piece processing mode, certainty, operability, and single objective), the research focus on scheduling has gradually shifted from the classical scheduling problem to the new scheduling problem. The semiconductor manufacturing system scheduling discussed in this book belongs to this new type of scheduling problem.

As for the process scheduling problem of semiconductor manufacturing system, the author paid attention to it as early as more than ten years ago. Based on the existing research foundation of flexible manufacturing system, discrete event dynamic system and Petri net, the author has tracked and understood the problems and current situation in this field and carried out some preliminary research. Until recent years, with the gradual deepening of the scientific research work of the author and research teams, they have received funding from a number of scientific research projects such as the National Key Basic Research Program of China (973 Program), the National Natural Science Foundation of China, enterprise cooperation, etc., and the research on the optimal scheduling of complex manufacturing processes has been further carried out systematically. The team has fully investigated and understood the theories and methods in this field, accumulated some research achievements, and cooperated with famous semiconductor manufacturing enterprises to make a beneficial exploration and attempt in the application and implementation of relevant research achievements. This book is written on the basis of these work and achievements, which combines the relevant research achievements in the recent 20 years and the author's years of accumulated experience and attempts to systematically discuss the intelligent scheduling problem of complex semiconductor manufacturing system from theory to method to application.

The intelligent scheduling problem of semiconductor manufacturing is a subject with NP characteristics and high complexity and challenge. With the continuous in-depth research on this kind of problem in the world and in response to the needs of the rise and rapid development of semiconductor manufacturing industry in China, the relevant research is still developing and improving. This book is only a phased summary of the author's study and research in recent years, which will inevitably have inadequacies; please criticize and correct.

Shanghai, China

Li Li  
Qingyun Yu  
Kuo-Yi Lin  
Yumin Ma  
Fei Qiao

# Contents

<b>1 Scheduling of Semiconductor Manufacturing System</b>	1
1.1 Semiconductor Manufacturing Process	1
1.2 Scheduling of Semiconductor Manufacturing System	6
1.2.1 Scheduling Characteristics	6
1.2.2 Scheduling Types	9
1.2.3 Scheduling Methods	12
1.2.4 Evaluation Indicators	14
1.3 Scheduling Development Trend of Semiconductor Manufacturing System	16
1.3.1 Data Preprocessing of Complex Manufacturing System	17
1.3.2 Data-Based Scheduling Modeling	18
1.3.3 Data-Based Scheduling Optimization	20
1.3.4 Analysis of Research Status	22
1.4 Summary	23
References	23
<b>2 Data-Driven Scheduling Framework of Semiconductor Manufacturing System</b>	25
2.1 Design of Data-Driven Scheduling Framework	25
2.2 Data-Based Scheduling Architecture of Complex Manufacturing System	28
2.2.1 Overview of DSACMS	28
2.2.2 Formal Description of DSCAMS	32
2.2.3 DSACMS-Based Modeling and Optimization of Scheduling for Complex Manufacturing Systems	42
2.2.4 Key Technologies in DSACMS	43
2.3 Application Examples	44
2.3.1 Overview of Fabsys	44
2.3.2 Object-Oriented Simulation Model of Fabsys (OOSM <sub>fab</sub> )	46



- 2.3.3 Data-Driven Forecasting Model in FabSys ..... 61
- 2.4 Summary ..... 64
- References ..... 64
- 3 Data Preprocessing of Semiconductor Manufacturing System ..... 67**
  - 3.1 Introduction ..... 67
  - 3.2 Data Standardization ..... 71
    - 3.2.1 Data Normalization Rules ..... 71
    - 3.2.2 Correction of Abnormal Values for Variables ..... 71
  - 3.3 Filling of Missing Data ..... 73
    - 3.3.1 Filling Method for Missing Data ..... 73
    - 3.3.2 Memetic Algorithm and Memetic Calculation ..... 75
    - 3.3.3 Attribute Weighted K Nearest Neighbor Missing Value Filling Method (KNN) Based on Gaussian Mutation and Depth First Search (GD-MPSO): GD-MPSO-KNN ..... 78
    - 3.3.4 Numerical Verification ..... 80
  - 3.4 Outlier Detection Based on Data Clustering Analysis ..... 82
    - 3.4.1 Outlier Detection Based on Data Clustering ..... 82
    - 3.4.2 K-Means Clustering ..... 84
    - 3.4.3 Data Clustering Algorithm Based on GS-MPSO and K-Means Clustering (GS-MPSO-KMEANS) ..... 85
    - 3.4.4 Numerical Verification ..... 86
  - 3.5 Redundant Variable Detection Based on Variable Clustering ..... 88
    - 3.5.1 Principal Component Analysis ..... 88
    - 3.5.2 Variable Clustering Based on K-Means Clustering and PCA ..... 90
    - 3.5.3 Variable Clustering Algorithm Based on MCLPSO (MCLPSO-KMEANSVAR) ..... 92
    - 3.5.4 Numerical Verification ..... 93
  - 3.6 Summary ..... 96
  - References ..... 96
- 4 Correlation Analysis of Performance Index of Semiconductor Production Line ..... 99**
  - 4.1 Long-Term and Short-Term Performance Indicators of Semiconductor Manufacturing System ..... 99
  - 4.2 Statistical Analysis of Performance Indicators for Semiconductor Production Lines ..... 104
    - 4.2.1 Short-Term Performance Indicators ..... 106
    - 4.2.2 Long-Term Performance Indicators ..... 110
  - 4.3 Correlation Analysis of Performance Indicators Based on Correlation Coefficient Methods ..... 112
    - 4.3.1 Block Diagram of Correlation Analysis ..... 113
    - 4.3.2 Correlation Analysis of Performance Indicators Considering Working Conditions ..... 114

- 4.3.3 Correlation Analysis of Performance Indicators Considering Dispatching Rules ..... 118
- 4.3.4 Correlation Analysis of Performance Indicators Considering Working Conditions and Dispatching Rules ..... 119
- 4.3.5 Correlation Analysis Between Long-Term Performances and Short-Term Performances ..... 119
- 4.3.6 Correlation Analysis of Performances on Production Line Named MIMAC ..... 120
- 4.4 Correlation Analysis of Performances Based on Pearson Coefficient ..... 124
  - 4.4.1 Correlation Analysis of Daily WIP and Daily Moving Steps ..... 125
  - 4.4.2 Correlation Analysis of Daily Queue Leader and Daily Moving Steps ..... 126
  - 4.4.3 Correlation Analysis of Daily Equipment Utilization and Daily Moving Steps ..... 128
- 4.5 Data Set of Performances for Semiconductor Manufacturing System ..... 129
  - 4.5.1 Training Set of Processing Cycle and Corresponding Short-Term Performances ..... 130
  - 4.5.2 Training Set of On-Time Delivery Rate and Corresponding Short-Term Performances ..... 131
  - 4.5.3 Training Set of Waiting Time and Corresponding Short-Term Performances ..... 132
- 4.6 Summary ..... 132
- References ..... 132
- 5 Data-Driven Release Control of Semiconductor Manufacturing System ..... 135**
  - 5.1 Common Release Strategies of Semiconductor Manufacturing Systems ..... 135
    - 5.1.1 Common Release Control ..... 136
    - 5.1.2 Improved Release Control Strategy ..... 145
    - 5.1.3 Research Status of Release Control Strategies ..... 147
  - 5.2 Release Control Strategy Based on Extreme Learning Machine ... 148
    - 5.2.1 Release Control Strategy for Determining Releasing Time Based on ELM ..... 150
    - 5.2.2 Release Control Strategy Based on Limit Learning Machine to Determine Releasing Sequence ..... 158
  - 5.3 Optimization of Release Control Based on Attribute Selection ... 167
    - 5.3.1 Attribute Set Related to Releasing ..... 168
    - 5.3.2 Attribute Selection ..... 168
    - 5.3.3 Simulation Based on Attribute Selection ..... 172

5.4	Summary .....	180
	References .....	181
<b>6</b>	<b>Data-Driving Dynamic Scheduling of Semiconductor Manufacturing System</b> .....	<b>183</b>
6.1	Dynamic Dispatching Rules .....	183
6.1.1	Definition of Parameters and Variables .....	184
6.1.2	Hypothesis .....	185
6.1.3	Decision-Making Process .....	185
6.1.4	Simulation and Verification .....	190
6.2	Optimization of Algorithm Parameters Based on Data Mining .....	192
6.2.1	Overall Design .....	192
6.2.2	Algorithm Design .....	194
6.2.3	Process of Optimization .....	202
6.2.4	Simulation and Verification .....	203
6.3	Summary .....	210
	References .....	210
<b>7</b>	<b>Performance-Driving Dynamic Scheduling of Semiconductor Manufacturing System</b> .....	<b>211</b>
7.1	Performance Prediction Method .....	211
7.1.1	Prediction Method of Long-Term Performances for Single-Bottleneck Semiconductor Production Line .....	211
7.1.2	Prediction Method of Long-Term Performances for Multi-bottleneck Semiconductor Production Line .....	221
7.2	Dynamic Scheduling of Semiconductor Production Line Based on Load Balancing .....	231
7.2.1	Overall Design .....	231
7.2.2	Load Balancing Technology .....	232
7.2.3	Selection of Parameters .....	233
7.2.4	Forecasting Model of Load Balancing .....	234
7.2.5	Dynamic Scheduling Algorithm Based on Load Balancing .....	237
7.2.6	Simulation and Verification .....	244
7.3	Dynamic Scheduling of Semiconductor Production Line Driven by Performances .....	245
7.3.1	Structure of Performance-Driving Scheduling Method .....	247
7.3.2	Dynamic Dispatching Rules .....	249
7.3.3	Prediction Model .....	252
7.3.4	Simulation and Verification .....	254
7.4	Summary .....	254
	References .....	255

- 8 Development Trend of Scheduling Problems for Semiconductor Manufacturing System Under Big-Data** ..... 257
  - 8.1 Industry 4.0 ..... 257
  - 8.2 Industrial Big Data ..... 260
  - 8.3 Development Trend of Scheduling Problems for Semiconductor Manufacturing Under Big-Data ..... 262
    - 8.3.1 Data-Based Petri Net ..... 262
    - 8.3.2 Dynamic Simulation ..... 262
    - 8.3.3 Prediction Model ..... 263
  - 8.4 Application Example: Big Data Driving Forecasting Model in Complex Manufacturing System ..... 263
  - 8.5 Summary ..... 265
- References ..... 266

# Chapter 1

## Scheduling of Semiconductor Manufacturing System



Semiconductor is the core of many industrial complete equipment, widely used in computer, consumer electronics, network communication, automotive, industrial/medical, military/government, and other core fields. With the popularity of the concept of “intelligence”, the importance of chip industry is becoming more and more significant. To get rid of the “pain of lack of core”, China has vigorously supported the domestic semiconductor industry from both the policy and capital aspects, striving for the realization of independent replacement. This chapter mainly introduces the scheduling of semiconductor manufacturing systems and its development trend, including the scheduling process, scheduling characteristics, scheduling types and scheduling methods, evaluation indexes, and scheduling problems.

### 1.1 Semiconductor Manufacturing Process

The semiconductor industry, as the most important part of the electronic components industry, is mainly composed of four components: integrated circuits (about 81%), optoelectronic devices (about 10%), discrete devices (about 6%), and sensors (about 3%). Therefore, semiconductor and integrated circuits are usually equivalent. Integrated circuits are divided into four main categories by product type: logic devices (about 27%), memory (about 23%), microprocessors (about 18%), and simulator components (about 13%). Semiconductor is a demand-driven market. Over the past four decades, the growth of the semiconductor industry has been driven by a shift from the traditional PC and related industries to the mobile market (including smartphones and tablets). In the future, it is likely to shift to wearables and VR/AR devices. From 2000 to 2015, the annual growth rate of China’s semiconductor market led the world, up to 21.4% (the annual growth rate of the global semiconductor market was

3.6%, of which the Asia–Pacific region was about 13%, the United States was nearly 5%, and Europe and Japan were all lower). In terms of global market share, China's semiconductor market share has increased from 5 to 50%, becoming the core market of the global semiconductor industry [1–3].

In 2015, the three major fields of integrated circuits showed a growth trend. The design industry saw the fastest growth, with sales of \$21.57 billion, up 26.55% year by year. Sales revenue of the chip manufacturing industry was \$14.67 billion, up 26.54% year by year; Package and test sales were \$22.52 billion, up 10.19% from the previous year. In terms of the proportion of industrial chain, the proportion of China's design industry is growing fastest, the proportion of packaging and testing has declined, and the proportion of manufacturing has remained stable. Benefiting from the policy support and the development of domestic economy, the three major IC structures tend to be optimized gradually: in 2015, the proportion of China's IC design accounted for 36.70%, manufacturing accounted for 24.95%, packaging and testing accounted for 38.34%; Chip sales amounted to 90.08 billion yuan, with an increase of 26.5%, 8 percentage points higher than the growth rate in 2014 [4–7].

The IC industry chain can be roughly divided into three major areas: circuit design, chip manufacturing, packaging, and testing. Integrated circuit production process is dominated by circuit design, which requires a variety of high-precision equipment and high-purity materials. Its general process is as follows: design companies provide integrated circuit design schemes, chip manufacturers produce wafers, packaging plants package and test integrated circuits, and sales to electronic product enterprises [8].

The semiconductor manufacturing process can be simply divided into wafer manufacturing and integrated circuit manufacturing. Among them, wafer manufacturing roughly includes several steps of purification of ordinary silica sand (quartz sand), molecular pulling–crystal cylinder (cylindrical crystal), and wafer (cutting the wafer into circular wafers) [9]. Among them, molecular crystallization means that the obtained high purity polycrystalline silicon is melted to form liquid silicon, and the silicon seed of single crystal contacts with the liquid surface and slowly pulls up while rotating. Finally, as the silicon atoms leave the liquid surface to solidify, a column of monocrystalline silicon is formed, which is as pure as 99.999999%. Cutting a wafer refers to cutting a silicon wafer with certain specifications from a single crystal silicon rod. These silicon wafers will be washed, polished, cleaned, inspected by human eyes and machine, and finally be inspected for surface defects and impurities by laser scanning. Qualified wafers will be delivered to chip manufacturers [10].

The IC fabrication process is the focus of this book. It consists of a variety of single processes and consists of three steps: thin film fabrication process, graphic transfer process, and doping process. The specific manufacturing process is shown in Fig. 1.1.

### (1) Film preparation technology

Thin film preparation means to grow several layers of films with different materials and thicknesses on the surface of the wafer, and the main processes include

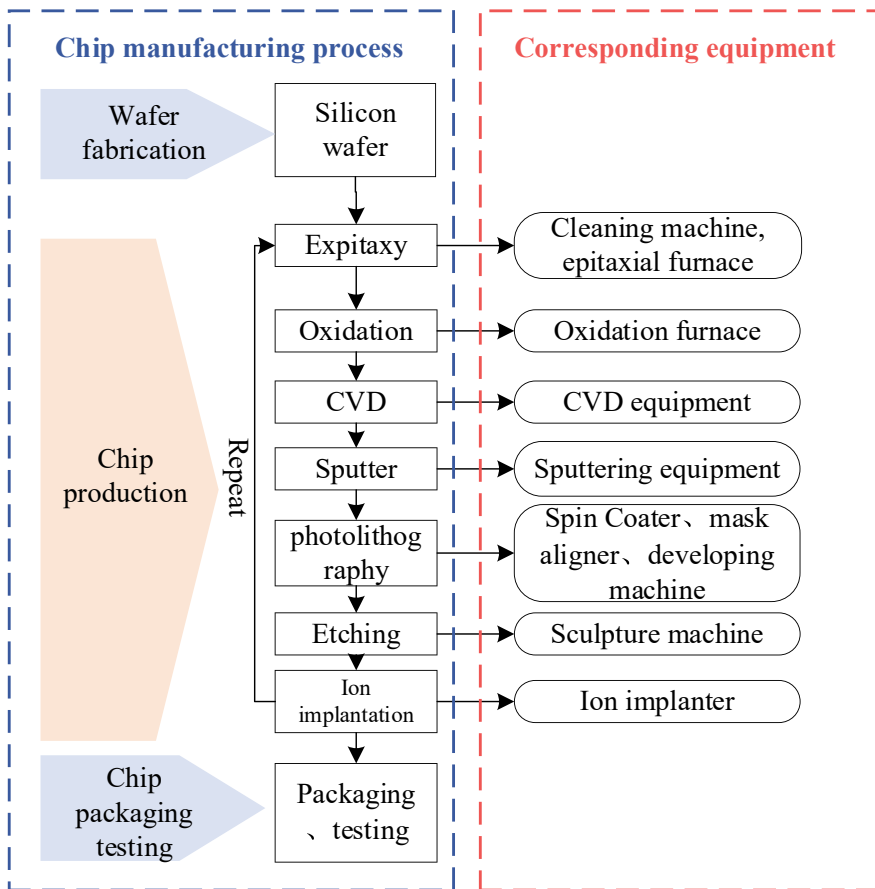


Fig. 1.1 Integrated circuit manufacturing process

three methods: oxidation, chemical vapor deposition (CVD) and physical meteorological deposition (PVD) [11].

- Oxidation: The reaction of a wafer with an oxygen-containing substance (oxygen or an oxidant such as water vapor) at high temperatures to form a thin film of silicon dioxide.
- CVD: one or several compounds or elemental gases containing the elements that constitute the film are passed into the reaction chamber with the substrate, and the solid film is deposited on the surface of the substrate by means of a space vapor phase chemical reaction.
- PVD: The material source is ionized into ions by physical method, and the film with some special function is deposited on the surface of the substrate through the action of low-pressure gas or plasma.

## (2) Graphic transfer process

The process of oxidation, deposition and diffusion, and ion implantation in the Integrated Circuit (IC) manufacturing process is not selective to the wafer. They are all processed on the whole silicon wafer without any graphics. The core of IC manufacturing is the transfer of the design pattern to a silicon wafer through a pattern transfer process (mainly a lithography process). As one of the most important process steps of semiconductors, lithography is to copy the graphics on the mask onto the silicon wafer. The cost of lithography is about 1/3 of the whole silicon wafer manufacturing process, and the time required accounts for about 40–60% of the whole silicon wafer manufacturing process. The process steps are as follows:

- The silicon wafer is coated with a photoresist and covered with a photoresist mask plate with a certain pattern made in advance.
- Exposure to the wafer coated with photoresist (the properties of photoresist will change after the photoresist is sensitized, the photosensitive part of the positive adhesive becomes easily dissolved, while the negative adhesive is the opposite).
- The development of the wafer (after the development of the gel is dissolved, leaving only the unilluminated part of the pattern; Negative glue, on the contrary, the part of the light is not easy to dissolve).
- The wafer is etched to remove the part that is not covered by the photoresist, and then the pattern on the photoresist is transferred to the underlying material.
- To remove photosensitive glue from a wafer by means of a degluing process.

## (3) Doping process

Doping is a process in which controlled amounts of impurities are added to specific areas of a wafer to alter the electrical properties of a semiconductor. Diffusion and ion implantation are two main processes of semiconductor doping [12].

- Diffusion: The process by which atoms, molecules, or ions are driven by high temperatures (900–1200 °C) from a high concentration zone to a low concentration zone. The concentration of impurity decreases monotonically from surface to body and the distribution of impurity is determined by temperature and diffusion time.
- Ion implantation: In a vacuum system, ions are accelerated by an electric field and changed by a magnetic field so that ions can be injected into the wafer with a certain amount of energy to form an injection layer with special properties in a fixed area and achieve the purpose of doping.



Compared with other manufacturing systems, semiconductor manufacturing systems have the following three distinct characteristics:

(1) Complicated technological process

Production technological process refers to the method and process in which workers use production tools to process all kinds of raw materials and semi-finished products through certain equipment and in a certain sequence, and finally make them into finished products. That is the combination of elements in the production process of products, from raw materials to finished products. The average processing cycle of a silicon wafer in the production line is relatively long, generally about 1 month. The process flow of silicon wafer varies from product to product, from dozens of steps to hundreds of steps, which also leads to the dispersed processing cycle of silicon wafer. A typical process is 250–600 steps, with 60–80 types of equipment used. In addition, there may be different orders and product categories on the production line. There are dozens of products produced on the production line, and there is a large number of re-entrant phenomena in the process, which will make the work-in-process competition for the right to use online equipment very fierce [13].

(2) Multiple entry process

In semiconductor manufacturing, reentrant is the nature of the system. Similar parts at different stages of processing may be processed simultaneously in front of the same equipment, and the parts may repeatedly access some equipment at different stages of the processing process. There are two main reasons for this: first, semiconductor components are hierarchical structures, each layer is produced in the same way, but the added material or accuracy is different; Second, semiconductor processing equipment is very expensive, the need to maximize the use of equipment, resulting in the emergence of multiple into the processing process. In short, the phenomenon of reentrant greatly increases the number of pieces to be processed for each device. In addition, the different types, quantities, and combinations of products, as well as the different complexity of the process flow of each product, make the scheduling and control of the semiconductor production line more complex [14].

(3) Mixed processing mode

Due to the different types of semiconductor production line equipment, its processing methods are also diversified. According to the processing mode of equipment, it is mainly divided into single chip processing, serial batch processing, single card parallel batch processing, and multi-card parallel batch processing. The existence of a hybrid machining mode further increases the complexity of semiconductor production line scheduling. At present, a large number of studies are based on simplified processing methods (single card processing and batch processing).

## 1.2 Scheduling of Semiconductor Manufacturing System

Production scheduling, as one of the effective ways to improve the economic benefit and market competitiveness of enterprises, is also a research hotspot in industrial engineering, management engineering, automation, and other fields. Generally speaking, production scheduling is to optimize the execution efficiency or cost of a decomposable production task by determining the processing sequence of the workpiece and the allocation of scheduling resources on the premise of satisfying the constraints of technology and resources. As a research proposition with a long history, the requirements of production scheduling include satisfying constraints, optimizing performance and practical efficiency. Its basic tasks can be summarized as modeling and optimization, that is, the understanding of scheduling problems and the solution of scheduling problems. Since the 1950s, scholars at home and abroad have carried out research and exploration on these two tasks, and some advanced scheduling modeling technologies and optimization methods have been put into practice and successfully applied [15].

### 1.2.1 Scheduling Characteristics

Semiconductor manufacturing system is different from the traditional production mode of job-shop and flow-shop. Whether it is a job shop or flow shop, different processes in the process flow need to be completed on different equipment. A significant feature of semiconductor manufacturing systems is that different processes of the workpiece may repeatedly access the same device, resulting in a large amount of reentry process flow. In the 1990s, Kumar defined semiconductor manufacturing system as the third type of production system developed after the job shop and flow shop—reentrant production system. With the complexity of integrated circuit performance and the miniaturization of component size, semiconductor manufacturing process has become more complex and sophisticated [16]. Therefore, semiconductor manufacturing is recognized as one of the most complex manufacturing systems at present, and its scheduling problem has also been widely concerned by academia and industry and has become a research hotspot in the field of control and industrial engineering.

From the perspective of the scheduling problem, there are two main modes of manufacturing system: Job-shop and Flow-shop. The semiconductor production line differs greatly from these two typical modes because of its obvious reentrant characteristics. The scheduling problem not only has the characteristics of general scheduling problems but also has some obvious special complexity:

#### (1) Non-zero initial state

As mentioned above, the average processing cycle of semiconductor silicon wafer is longer, generally dozens of days. During this period, new wafers will

be put into the production line every day for processing, rather than waiting for all wafers on the production line to be processed and then put in new wafers. Therefore, when the semiconductor production line is initially scheduled, there are already a large number of wafers in the production line, some in the state of being processed, and some in the state of being processed. This non-zero initial state is an important characteristic of scheduling problems in semiconductor manufacturing systems.

(2) On a large scale

The semiconductor production line consists of hundreds of devices, each product process includes hundreds of processes, and there may be as many as hundreds of product types flowing on the production line. This makes the scheduling problem of semiconductor production line much larger than the typical scheduling problem, which makes the complexity of the problem and the difficulty of solving greatly increased. In the production process, the workpieces, equipment, operators, logistics transmission system and buffer zone in the workshop are mutually influenced and restricted [17]. Therefore, not only the loading and processing time of each workpiece, the number of equipment, the buffer capacity of the system, the processing sequence of the workpiece and other resource factors should be considered, but also the uncertain factors such as the operation proficiency of personnel, and even the influence of various dynamic events on scheduling should be considered. Therefore, the scheduling problem of job shop is actually a complex combinatorial optimization problem with many constraints. With the increase of scheduling scale, the amount of computation needed to obtain the feasible solution increases exponentially, and the possibility of obtaining the optimal solution or near-optimal solution becomes less and less.

(3) Uncertainty

The uncertainty of semiconductor production line scheduling is mainly manifested in the following three aspects:

- ① The total number of tasks is uncertain: in the actual semiconductor production line, new wafers are constantly put into processing every day, but only the number of new wafers invested in a period of time (such as a day or three days) can be known, and the total number of wafers is uncertain.
- ② Uncertain events: the existence of a large number of uncertain events in the semiconductor manufacturing system usually causes a change in the state of the production line; Therefore, it is necessary to consider the influence of these uncertain events in the scheduling of semiconductor manufacturing systems.
- ③ Process processing time is uncertain: on the one hand, the same process in different equipment for processing, the required processing time may be different, and with the aging of some parts, the same process on the same

equipment processing time will also produce greater changes; On the other hand, some processes need to test pieces (i.e., trial processing) before formal processing, and the test time may change with the change of the number of test pieces, resulting in the uncertain processing time of a card of silicon wafer in this process [18].

(4) Scheduling scheme has short validity

In addition to the current WIP situation on the production line, the formulation of scheduling scheme should also have a detailed and determined releasing plan (determined product type and quantity). Although a certain number of new wafers are put into the production line for processing every day, generally, only a short period of time (such as one day or three days) can be determined for the releasing plan. Therefore, compared with the average production cycle of semiconductor products ranging from ten days to dozens of days, the effective period of the actual semiconductor production line scheduling scheme is generally short. Coupled with the occurrence of a large number of uncertain events, the validity period of the scheduling scheme is difficult to exceed one day.

(5) Local optimization problem

In shop floor production, there are many different production tasks, which may have different requirements for scheduling goals, and sometimes these requirements are mutually exclusive. For example, the requirements of small production cycle, the least overdue orders, the highest equipment utilization rate, and so on. Therefore, how to make the production scheduling system meet these goals as much as possible is also a constant problem for workshop production scheduling [19]. Due to the short validity period of the scheduling scheme, the optimization of the production line system performance index can only be short-term and partial, and only part of the system performance index can be optimized, such as equipment utilization rate, total movement volume, movement rate, etc. However, indicators such as average processing cycle and its variance, just-in-time delivery rate, and delay rate cannot be significantly optimized.

(6) Binding

Constraints are mainly reflected in two aspects: process path constraint and resource constraint. First of all, each product, whether simple or complex, has strict process path constraints, and usually, the sequence of each process cannot be reversed. Secondly, the supply of processing raw materials, the scale of production equipment, the production capacity of production equipment, etc., are not infinite. Therefore, production scheduling is carried out under multiple constraints.

In general, semiconductor manufacturing system scheduling has obvious characteristics of multiple entry, high uncertainty of manufacturing environment, high complexity of manufacturing process and multi-objective optimization of scheduling objectives. Accordingly, dynamic scheduling methods that can respond to real-time operating environment have been paid more attention.

### 1.2.2 Scheduling Types

- **Scheduling classification of semiconductor production line based on scheduling object**

Semiconductor manufacturing production lines have large scales and different types of equipment, which can be divided into workpiece scheduling, feed control, bottleneck scheduling, batch processing equipment scheduling, production line scheduling, and maintenance scheduling according to different concerns.

(1) Workpiece scheduling

Workpiece scheduling strategy has a direct impact on the performance of the production system, so workpiece scheduling is the focus of semiconductor manufacturing system scheduling research. There are five kinds of scheduling methods: traditional operations research, discrete system simulation, mathematical model, computational intelligence, and artificial intelligence.

(2) Release control

Release control is to determine when and how much raw materials are put into the production system under the guidance of a certain releasing strategy so as to maximize the production capacity of the production system, generally divided into static releasing and dynamic releasing two ways. Static releasing is based on the pre-set rate (such as fixed time interval releasing or random distribution of Poisson flow releasing), because the actual changes of the production line can not be tracked, easy to cause workpiece backlog, so that the performance of the production line decline; Dynamic releasing is based on the actual situation of the production line (such as delivery time and WIP level and other performance indicators), using a heuristic method to control releasing.

(3) Bottleneck equipment scheduling

Bottleneck equipment scheduling refers to the optimal scheduling scheme of the entire semiconductor production line by solving the scheduling problem of the bottleneck area. The loss of capacity at the bottleneck is the loss of the whole plant, so the scheduling problem of bottleneck equipment is very important.

Bottleneck can be identified by observing the queue length of WIP and measuring the utilization rate of machine. ① Analyze the queue length of WIP in the manufacturing system. In this method, either queue length

or wait time is measured, and the machine with the longest queue length or wait time is considered to be the manufacturing bottleneck. The advantage of this method is that the instantaneous manufacturing bottleneck of the system can be detected by a simple comparison of queue length or waiting time. The disadvantage is that many production systems have limited WIP queues or none at all, in which case the queue length method cannot be used to detect manufacturing bottlenecks. ② Measuring the utilization rate of different machines in the production system, the machine with the highest utilization rate is the manufacturing bottleneck. However, the utilization rates of different machines are often very similar, so this method cannot say with certainty which machine is the manufacturing bottleneck. Long-time simulations may yield accurate and meaningful results, but this approach is limited by steady-state systems. The method of measuring utilization can not determine the instantaneous manufacturing bottleneck, but only the average bottleneck over a long period of time. It is not appropriate to use it to detect and monitor the transfer of bottleneck.

(4) Batch processing equipment scheduling

Batch processing equipment scheduling is an important part of semiconductor production line scheduling, which has an important effect on the performance of semiconductor production line. Batch processing is the distinguishing feature of semiconductor manufacturing system. The multi-card parallel batch processing equipment, such as oxidized furnace tube, accounts for about 20–30% of the total equipment in the semiconductor manufacturing line, and its scheduling scheme is of great significance in improving the performance of semiconductor manufacturing system.

(5) Production line scheduling

Production line Scheduling pays attention to the flow direction of the workpiece in the whole semiconductor production line and its processing sequence on each equipment, and arranges the processing sequence and start processing time of the workpiece on each processing equipment, namely processing Scheduling, Dispatching, and Sequencing. There are many research methods used in the scheduling of semiconductor manufacturing systems, including traditional operations research, discrete event simulation technology and heuristic rules, as well as advanced artificial intelligence, computational intelligence, and swarm intelligence algorithms.

(6) Equipment maintenance scheduling

Equipment maintenance scheduling is a process used to determine when to remove equipment from the production line for scheduled maintenance. The equipment Maintenance mainly includes Preventive Maintenance (PM) and Corrective Maintenance (CM). The goal of preventive maintenance is to seek a compromise between planned and unplanned downtime. Corrective maintenance refers to the corresponding maintenance of equipment after an unexpected failure. The latter, due to unexpected equipment

failure, will lead to higher costs. At present, research on equipment maintenance scheduling mainly focuses on operations research methods or heuristic rules.

- **Scheduling classification of semiconductor production line based on scheduling environment and tasks**

Based on the different scheduling environments and tasks, the scheduling of semiconductor manufacturing systems can be divided into static scheduling and dynamic scheduling.

- (1) Static scheduling

Static scheduling refers to the process of forming an optimal scheduling scheme under the premise that the state of the manufacturing system and the processing task are determined. Static scheduling takes place at some point in time  $t_0$ . Of the manufacturing system state  $U(T_0)$ , determined workpiece information (specific processing task description) and time length  $T_0$  (commonly referred to as scheduling depth) is the input, and the appropriate scheduling algorithm is adopted to generate the scheduling cycle under the condition of satisfying the constraint conditions and optimization objectives  $[t_0, t_0 + T_0]$  in the scheduling scheme. The constraint conditions of static scheduling include system resources, product process flow, delivery time, etc., and the optimization objective includes evaluation of the processing cycle, delivery time, performance indicators of manufacturing system, such as equipment performance rate, productivity, etc. Once the scheduling scheme is produced, the processing scheme of all the workpieces is determined and will not be changed in the later processing process.

- (2) Dynamic scheduling

Dynamic scheduling refers to the process of dynamically generating a scheduling scheme according to the state of manufacturing system and the actual situation of machining task. There are two ways to realize dynamic scheduling: one is to adjust the static scheduling scheme in time and generate a new scheduling scheme based on the existing static scheduling scheme and the field state and processing task information of the manufacturing system. This scheduling process is also called rescheduling; Second, there is no static scheduling scheme in advance, and the processing task is determined for idle equipment directly according to the real-time state of the manufacturing system and the processing task information. This scheduling process is also called real-time scheduling.

Both of the above two methods can obtain a highly operable scheduling scheme, but the optimization calculation process is different. The real-time scheduling usually only considers the local information in the decision-making, so the scheduling scheme obtained is only feasible, which may have a large distance from the optimal scheduling scheme. Rescheduling is based

on the existing static scheduling scheme. According to more system state information and processing task information to dynamically adjust the static scheduling scheme, the obtained scheduling scheme is not only operable, but the optimization effect is also better, closer to the optimal scheduling scheme.

Compared with static scheduling, dynamic scheduling can produce a more operable decision-making scheme according to the actual situation of production site. In view of the characteristics of dynamic scheduling, the following two factors must be fully considered: first, the optimization process must make full use of the real-time information reflecting the state of manufacturing system and processing task; second, the dynamic scheduling scheme must be completed within a short time without affecting the operation of equipment.

### ***1.2.3 Scheduling Methods***

At present, various scheduling methods can be roughly classified into five categories: the methods based on operations research, the methods based on one-step heuristic rules, and the methods based on artificial intelligence, computational intelligence, and swarm intelligence.

#### **(1) Methods based on operations research**

In this method, the production scheduling problem is transformed into a mathematical programming model, and the branch-and-bound method based on an enumeration idea or dynamic programming algorithm is used to solve the optimal or approximate optimal solution of the scheduling problem, which is an accurate algorithm. For complex problems, especially in the semiconductor wafer manufacturing industry with different production characteristics from the traditional Job-shop and Flow-shop, this pure mathematical method has the weaknesses of difficult model extraction, a large amount of computation and difficult algorithm realization. For the dynamic scheduling in the production environment, it is complicated to realize and cannot solve the dynamic and fast response problems.

#### **(2) Methods based on one-step heuristic rules**

Heuristic rule refers to the method of choosing some or some attributes of the workpiece as the priority of the workpiece and selecting the workpiece for processing according to the priority level. According to the different scheduling objectives, the semiconductor manufacturing process heuristic rules can be divided into the rules based on the delivery date, the rules based on the processing cycle, the rules based on the waiting time of the workpiece, the rules based



on whether the application procedure of the workpiece is the same, and the rules based on the load balancing. Heuristic rules have become the first choice for dynamic scheduling in the actual semiconductor manufacturing environment due to their simplicity and rapidness, but they also have some limitations, which can only improve the individual performance index of the product but have a weak ability to improve the overall performance of the production line.

Due to scheduling optimization of semiconductor manufacturing process is a very complicated problem. Its performance is good or bad depends not only on the scheduling policy itself, and the variance of the system model, processing time, the actual average processing cycle, and processing cycle than the relevant theory, and the system bottleneck in the device number, need to repeat visits, emergency orders to join factors have very close ties. Although heuristic rules require little computation, have high efficiency and good real-time performance, they usually only provide feasible solutions for one or more targets, and lack effective grasp and prediction ability for the overall performance. The scheduling results of heuristic rules may deviate from the global optimization of the system to a certain extent, or even a large deviation. Therefore, heuristic rules often need to be used in conjunction with intelligent methods that select among alternative rules based on the state of the system. Typical research methods usually use a combination of some kind of intelligent method, simulation method, and heuristic rules.

- (3) Methods based on artificial intelligence, computational intelligence, and swarm intelligence

Artificial intelligence, also known as machine intelligence, is a comprehensive discipline developed from the interpenetration of computer science, cybernetics, information theory, neurophysiology, psychology, linguistics, and other disciplines. The commonly used artificial intelligence systems in semiconductor scheduling algorithms include expert systems and artificial neural networks, among which the artificial neural network is usually combined with other methods (such as dynamic programming).

Computational intelligence is based on the behavioral patterns of human beings and organisms or the motion patterns of substances. It establishes algorithm models through mathematical abstraction and solves combinatorial optimization problems through computer calculation. The commonly used computational intelligence is tabu search, simulated annealing, genetic algorithm, artificial immune algorithm and so on. In semiconductor manufacturing system scheduling, it is possible to solve scheduling problems by either using a single computational intelligence method, combining different computational intelligence algorithms or combining computational intelligence algorithms with modeling techniques to obtain better performance.

Swarm intelligence is an algorithm and model that is inspired by the swarm behavior of social organisms and simulates the abstraction. Without centralized control and a global model, swarm intelligence provides the basis for finding solutions to complex distributed problems. Common swarm intelligence has an ant colony optimization algorithm, pheromone algorithm, particle swarm optimization algorithm, and so on. The application of swarm intelligence in semiconductor manufacturing system scheduling is relatively rare.

### ***1.2.4 Evaluation Indicators***

The purpose of scheduling semiconductor production line is to optimize its system performance. Combined with the characteristics of semiconductor production line, the main indicators used to measure the impact of scheduling results on the performance of semiconductor production line system include:

(1) Yield

The percentage of the total product that is acceptable, often referred to as the percentage of acceptable cores on a silicon wafer. The yield has a significant impact on the economic benefits of semiconductor production lines. Obviously, the higher the yield, the higher the economic benefits. The production rate is greatly affected by the equipment process, and the main effect of scheduling is to shorten the residence time of the workpiece in the workshop as far as possible, reduce the chance of chip contamination, to ensure a higher yield.

(2) Quantity of Work in Process (WIP)

The total number of uncompleted pieces of work in the production line, i.e., the total number of silicon wafers on the production line or the total number of wafers. Minimizing the WIP number is an optimization objective related to minimizing processing cycles. Try to control the WIP number of semiconductor production line to be equal to the expected value, which is related to the processing capacity of semiconductor production line. Even if the WIP number continues to decline below the WIP expectation, the processing cycle will not be significantly shortened; When the WIP quantity is higher than the desired target value, the more WIP quantity, the longer the processing cycle will be. In addition, the higher the number of WIP, the more capital occupied, which will directly affect the economic benefits of enterprises.

(3) Equipment Utility (Machine Utility)

The ratio of the time spent in the processing state to the time spent on the machine can be measured by the idle cost of the equipment. Equipment utilization is related to the number of WIPs. Generally speaking, the number of WIPs is large and the equipment utilization rate is high. However, when the number of WIPs is saturated, the device utilization will not improve even if the number of WIPs

increases again. Obviously, the higher the utilization of equipment, the greater the number of processed workpiece, and the greater the value created.

(4) Average processing cycle and its variance

Processing cycle refers to the time taken by a silicon wafer from entering the semiconductor production line to finishing the processing of all procedures and leaving the production line, also known as the wafer flow time. Average processing cycle refers to the average processing time of a multi-card silicon wafer in the same process, and its variance refers to the mean root square of the processing cycle of each card silicon wafer and its average processing cycle. The mean cycle time and its variance indicate the responsiveness and just-in-time delivery of the system.

(5) Total movement

A silicon wafer that has completed one step of processing is called moving one step. The total Movement is the total number of moves (cards · steps) of all silicon wafers in a unit of time (such as a shift, 12 h). The higher the total movement, the higher the number of processing tasks completed by the production line. Movement is an important index to measure the performance of semiconductor production line. The higher its value is, the higher the processing capacity of semiconductor production line is, and the higher the utilization rate of equipment is.

(6) Movement speed

The average number of moves (steps/card) of a silicon wafer per unit time (e.g., 12 h per shift). The higher the moving rate is, the faster the wafer flows in the production line and the shorter the average processing cycle is.

(7) Productivity

The number of cards or silicon wafers flowing out of the production line per unit time (usually shift or day). Ideally, productivity is equal to the rate of feed. It becomes inversely proportional to processing cycle, namely processing cycle is shorter, and criterion productivity is higher. The productivity of a semiconductor production line determines the cost of the final product, the cycle time, and customer satisfaction. Obviously, the higher the productivity, the higher the value created per unit of time, and the higher the processing efficiency of the production line.

(8) On time delivery rate

Percentage of the number of parts delivered on time (on time or ahead of time) as a percentage of the total number of parts completed.

(9) Tardiness

The number of undelivered parts as a percentage of the total number of finished parts. It is obvious that the just-in-time delivery rate is directly or indirectly related to the rate of finished product, productivity, processing cycle, WIP, and